

Spatial Hearing in Echoic Environments: The Role of the Envelope in Owls

Brian S. Nelson^{1,*} and Terry T. Takahashi¹

¹Institute of Neuroscience, University of Oregon, Eugene, OR 97403, USA

*Correspondence: bsnelson@uoregon.edu

DOI 10.1016/j.neuron.2010.07.014

SUMMARY

In the precedence effect, sounds emanating directly from the source are localized preferentially over their reflections. Although most studies have focused on the delay between the onset of a sound and its echo, humans still experience the precedence effect when this onset delay is removed. We tested in barn owls the hypothesis that an ongoing delay, equivalent to the onset delay, is discernible from the envelope features of amplitude-modulated stimuli and may be sufficient to evoke this effect. With sound pairs having only envelope cues, owls localized direct sounds preferentially, and neurons in their auditory space-maps discharged more vigorously to them, but only if the sounds were amplitude modulated. Under conditions that yielded the precedence effect, acoustical features known to evoke neuronal discharges were more abundant in the envelopes of the direct sounds than of the echoes, suggesting that specialized neural mechanisms for echo suppression were unnecessary.

INTRODUCTION

In nature, sounds from an object of interest reflect off of nearby surfaces and arrive at our ears shortly after the arrival of the direct sound. If the sound is long and the delay is short, the direct and reflected waveforms will summate in each ear, altering the spectral and temporal features of the inputs to the auditory system. Despite this complexity in the physical signal, we effortlessly localize the actual sound source and not the reflective surface. The perceptual dominance of the direct sound over the reflection, or echo, has been termed “localization dominance” (Litovsky et al., 1999) and has been documented in a variety of classes including birds, mammals, and invertebrates (Dent and Dooling, 2004). Localization dominance and its complement, lag “discrimination suppression”, which refers to the weaker perceptual salience of the echo, are components of the precedence effect (Blauert, 1997; Haas, 1951; Litovsky et al., 1999; Wallach et al., 1949). As the delay between the direct (lead) sound and the reflection (lag sound) increases, the direct sound’s dominance subsides, the reflection’s salience

increases, and the “echo threshold” is said to have been crossed.

Normally, a direct sound arrives slightly before its echo, and this time advantage has been the focus of most precedence effect studies, especially, those that use transient stimuli such as clicks. Click-like stimuli have the practical advantage of eliminating the temporal overlap of the leading and lagging sounds. In nature, however, sounds are often longer than the delay, so that the leading and lagging sounds overlap. Moreover, the onsets of naturalistic sounds are sometimes gradual, which can obscure the lead/lag relationship at the sound pair’s beginning. Interestingly, Zurek (1980) and, more recently, Dizon and Colburn (2006) showed that human listeners experience localization dominance even when the onset (and offset) time-of-arrival differences are removed from lead/lag sound pairs. Here, we test the hypothesis that the delays between corresponding envelope peaks, or “ongoing echo delays” (OEDs), may be sufficient to cause localization dominance.

This hypothesis predicts that localization dominance with lead/lag pairs having synchronized onsets (and offsets) would depend on the depth of envelope modulations. We test this prediction and examine possible neural correlates in the barn owl (*Tyto alba*), a nocturnal predator whose sound localization is guided by activity on a topographic map of auditory space generated in the external nucleus of its inferior colliculus (ICx; Knudsen and Konishi, 1978). Lesions of this auditory space map lead to scotoma-like deficits in sound localization (Wagner, 1993), and the resolving power of its neurons can account for the owl’s auditory spatial discrimination (Bala et al., 2003, 2007). Moreover, studies have shown that at short delays, the neural representation of the direct sound is stronger on the map than that of the echo (Keller and Takahashi, 1996b; Nelson and Takahashi, 2008; Spitzer et al., 2004). Correspondingly, owls localize the direct sound preferentially (Keller and Takahashi, 1996b; Nelson and Takahashi, 2008; Spitzer and Takahashi, 2006), and their ability to discriminate changes in the locations of simulated echoes is poorer (Spitzer et al., 2003).

We first confirm that owls, like humans, localize the leading sound preferentially with only OEDs. This preference, however, required deep amplitude modulations. With stimuli used in the behavioral experiments, we also demonstrate that responses on the auditory space map are considerably stronger to the leading sound, but again, only if the envelopes were deeply amplitude modulated. We propose a general mechanism for localization dominance based on temporal superposition and the temporal response properties of space-map neurons.

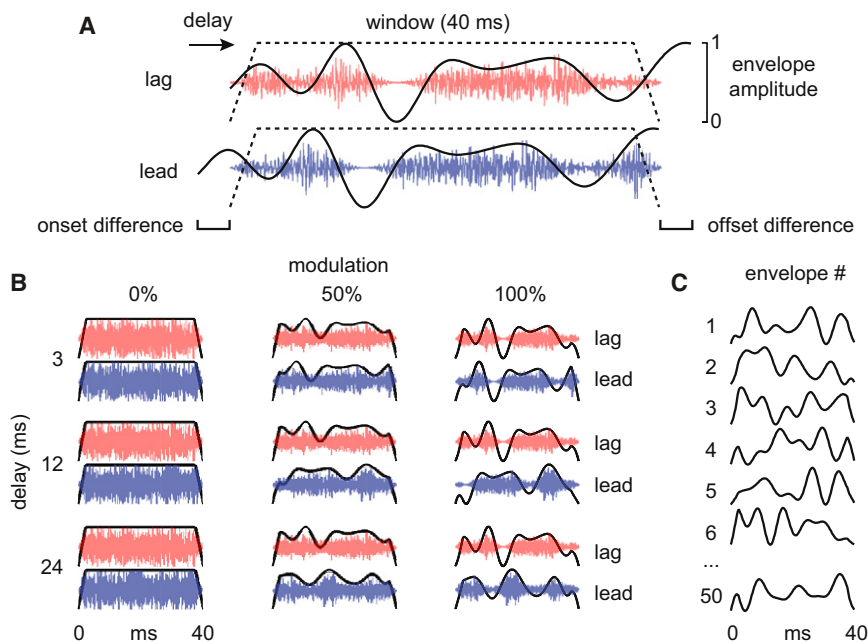


Figure 1. Stimulus Configurations

(A) Lead (bottom, blue) and lag (top, red) waveforms (100% modulation). Carriers were one of 50 arbitrary broadband noises (2–11 kHz). The stimuli were generated by delaying an identical copy of the lead sound 3, 12, or 24 ms and then truncating them so that the lead and lag sounds had synchronized onsets and offsets (trapezoidal dashed lines; 40 ms duration; 2.5 ms ramps). The delay was thus defined only by the ongoing echo delay (OED).

(B) Three envelope modulation depths and three OEDs resulted in nine different stimulus combinations. Note that 0% modulation (left column) indicates that no modulation was purposely imposed; only the noise carrier's intrinsic modulations were present.

(C) Examples of the arbitrary envelopes used. All envelopes had random phase spectra and amplitude spectra that spanned the owl's modulation transfer function (<150 Hz). For each behavioral trial, an envelope and carrier were chosen at random and applied to the lead and lag sounds. For physiology, all 50 envelopes were presented in a random order over one to three repetitions.

This mechanism requires no specialized circuitry for the “suppression” of echoes, such as lateral inhibition (Burger and Pollak, 2001; Fitzpatrick et al., 1995; Keller and Takahashi, 1996b; Lindemann, 1986; Pecka et al., 2007; Spitzer et al., 2004; Yin, 1994).

RESULTS

Behavior

We first determined whether or not owls experience localization dominance when only OEDs are available, by measuring the accuracy and precision of the owl's head turns. For owls, whose eyes are nearly immobile in their sockets, these head saccades are natural orienting responses analogous to the ocular saccades of primates. The stimuli, synthesized as explained in Figure 1, consisted of broadband noise bursts (2–11 kHz) multiplied by one of 50 different noise envelopes having spectra that encompassed the behavioral and neural (IC) modulation transfer functions of the barn owl (<150 Hz; Dent et al., 2002; Keller and Takahashi, 2000). The noise envelopes were modulated at 0%, 50%, or 100% before delaying the lag sound by 3, 12, or 24 ms and truncating the leading and lagging segments.

Prior to truncation, the lagging sound was an identical copy of the leading sound (same noise carrier, same envelope), but with a delay. For a trial, two speakers, diametrically opposed across the owl's center of gaze, were chosen to emit a direct sound and a simulated echo (Figure 2A). By thus positioning the sources, we could readily classify a saccade as being directed toward the lead or lag, even in cases when the error was large (Figure 2B). Sounds were also presented from single speakers to monitor each bird's localization performance in a standard localization task. We report data from three birds: C, S, and T.

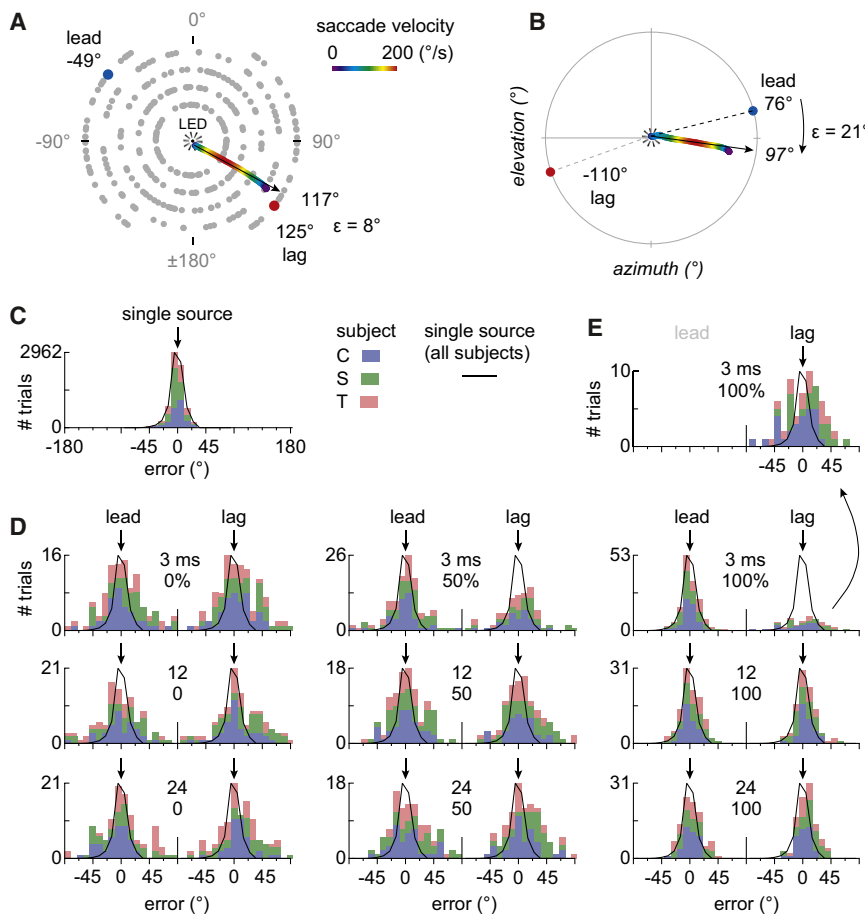
Figure 2C shows the distribution of head aims for single-source trials in all three birds. The abscissa represents localiza-

tion error (ϵ), measured from the start of each saccade, using polar coordinates, and in reference to the polar angle of the speaker to which each saccade was directed (Figure 2B). The center of each distribution represents saccade accuracy, and its breadth represents saccade precision. Saccade radii, shown in Figure S1, were highly variable across trials and did not vary consistently with delay or modulation depth.

The upper-left panel of Figure 2D shows the distribution of saccades elicited by lead/lag pairs with an OED of 3 ms and a modulation of 0%, i.e., when only the rather shallow modulations intrinsic to the noise carrier were present. Because the speaker pair was diametrically opposed across the central LED and the center of the owl's gaze, the lead and lag sources are shown to be 180° apart. The distribution is bimodal and symmetric, indicating that the owls localized the lead and lag sources with equal frequency. In earlier studies with similar stimuli, but with the delays at the onset and offset intact, owls showed clear preferences for leading stimuli (Nelson and Takahashi, 2008; Spitzer and Takahashi, 2006). Without the explicit delay at the onset (or offset), we now observed no evidence for localization dominance (0% modulation). The center and right panels in the upper row show results obtained at the same 3 ms delay but with deeper modulations (50% center; 100% right). These distributions are asymmetric, with more saccades to the leading source, suggesting that the owls experience localization dominance.

Thus, owls behaved in a manner consistent with localization dominance with stimulus pairs lacking onset (and offset) delays. Consistent with our hypothesis, the results also show that the ongoing delays in the *envelope* features were necessary and sufficient to induce localization dominance in owls.

Note that without the modulations (0%), the owls localized either source, suggesting that each sound of the pair was perceived as a spatially separate event. Although the two

**Figure 2. Head Saccades**

(A) Speaker locations used in head-saccade experiments. Gray dots mark the possible speaker locations. At the start of a trial, the owl was required to fixate an LED (center) before a speaker or speaker pair was activated. In trials with lead/lag pairs, the two speakers were diametrically opposed across the center of the array. The multi-colored lines in (A) and (B) show sample saccades and their trajectories (arrow). Colors represent the angular velocity of the saccade.

(C) Distribution of head-aims relative to single sources (0°) in three owls (C, S, and T). Each color represents a different bird. The black line is the normalized average distribution of head-aims for all three subjects.

(D) Distributions of head saccades in precedence-effect trials for OEDs of 3, 12, and 24 ms and modulations of 0%, 50%, and 100%. The locations of the lead and lag sources are indicated by arrows (0°). The distributions converge at the center of the abscissa because the speakers were separated by 180°. The distribution of lag-directed saccades for an OED of 3 ms is magnified and replotted in (E).

lead source (100%; subject C: $p = 0.046$, S: $p = 0.039$, T: $p = 0.0016$, all subjects: $p < 1 \times 10^{-6}$, contingency table analyses).

Inspection of Figure 2 shows that localization error was also affected by OED and modulation depth. The outline of the distribution of saccades to single sources (Figure 2C) is superimposed on each

histogram in Figure 2D and demonstrates that the distributions are broader (i.e., the saccades are less precise) than those toward single sources when modulations are shallow (Figure 2D, left and center columns). When the modulations were maximal (right column), saccade precision to the lead and lag sources rivals that to single sources.

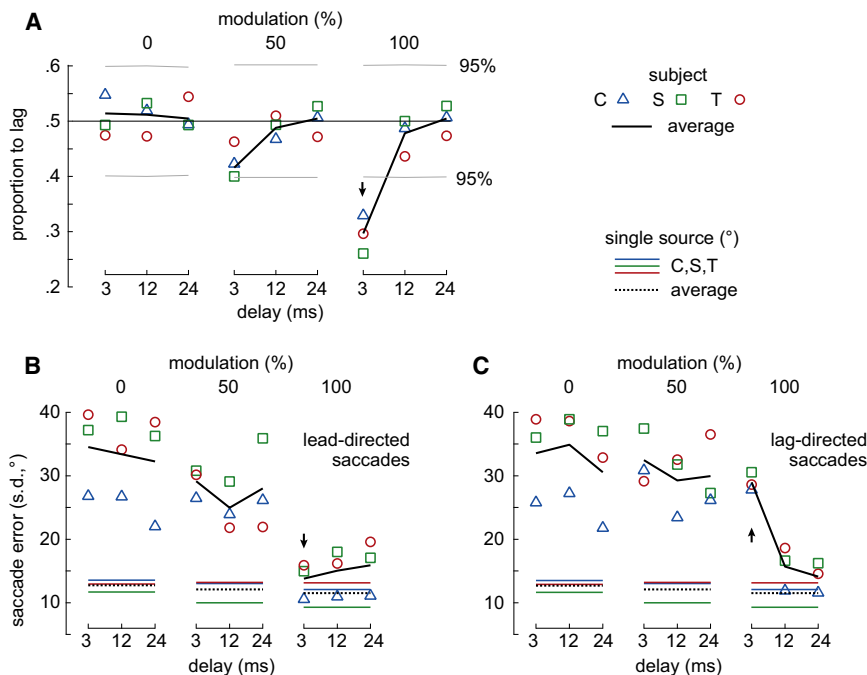
The trend noted above is more readily apparent in Figures 3B and 3C, which plot saccade error against delay for the three modulation depths. Comparing across panels of Figure 3B, one can see that lead-directed saccades increased in precision (i.e., decreased in error) with modulation depth and were nearly as precise as saccades to single sources (colored lines, black dashed line) when the modulation depth was 100% ($SD < \sim 20^\circ$).

Lag-directed saccades (Figure 3C) were generally imprecise for all OEDs at modulation depths of 0% and 50% and for the shortest OED (3 ms) at 100%. In the latter case, saccade precision was roughly equivalent to that toward unmodulated stimuli (0%; $SD > \sim 21^\circ$). For the longer delays at 100% modulation, saccades were nearly as precise as those to single sources. Comparison of points indicated by arrows in the right-hand panels of Figures 3B and 3C (at 3 ms), show that saccades to deeply modulated lag sounds were significantly less precise than those that were lead directed (subject C: $p = 0.0001$, S: $p = 0.0001$, T: $p = 0.0045$, all subjects: $p = 0.0001$; Levene's test).

The lower two rows in Figure 2D show results for longer OEDs (12, 24 ms). The distributions are all bimodal and symmetric, suggesting that localization dominance no longer operates at these longer OEDs, regardless of the modulation depth.

The effects of OED and modulation depth on localization dominance (Figure 2D) are summarized in Figure 3A, which plots the proportion of lag-directed saccades for all subjects (black line; $\pm 95\%$ confidence interval; gray horizontal lines) against OED for envelope modulations of 0% (left), 50% (center), and 100% (right). These data show that when the envelope modulations are minimal (left), owls saccade to the lag source on roughly half the trials, regardless of the OED. For deeper modulations and short OEDs (top, left), the owls preferentially localized the

histogram in Figure 2D and demonstrates that the distributions are broader (i.e., the saccades are less precise) than those toward single sources when modulations are shallow (Figure 2D, left and center columns). When the modulations were maximal (right column), saccade precision to the lead and lag sources rivals that to single sources.

**Figure 3. Head Saccade Characteristics**

(A) Proportion of lag-directed saccades for each bird (colored symbols). Black lines show the average of the birds. Horizontal gray lines and an arrow indicate proportions that differed significantly from 0.5 (log likelihood ratio, 95% confidence).

(B and C) Saccade error (ϵ), defined as the standard deviation (SD) of each saccade distribution (Figures 2C–2E), plotted against OED and modulation depth for lead-directed (B) and lag-directed (C) saccades. Colored lines show the error with single sounds.

See also Figure S1.

Physiology

The behavioral results above suggest that owls experience localization dominance with stimuli in which the lead/lag relationship is defined only by the ongoing envelope features of the two sounds. We now describe the responses evoked in neurons of the owl's auditory space map by stimuli used in the behavioral experiments.

The space-map neurons of our sample all had well-defined spatial receptive fields (SRF; $n = 28$ cells) as shown in Figure 4A, which represents the firing rate (color scale) of a typical cell evoked by stimuli placed at different locations in virtual auditory space. In the diagram, negative azimuths and elevations correspond to loci that are, respectively, to the bird's left and below its eye level. For trials with lead/lag pairs, one source, called the "target," was placed in the cell's SRF, and the other source, called the "masker," was placed outside the SRF at a mirror-symmetric location across the intersection of the owl's midsagittal plane and eye level (gray crosshair, Figure 4A). For any given trial, the sound source in the cell's SRF, i.e., the target, could lead or lag.

We also conducted trials with single sources at the locations of the target and masker, and with two sources simultaneously emitting noise bursts with *uncorrelated* carriers but identical, synchronous envelopes from the masker and target loci (no OEDs). The latter stimulus, which we will refer to as "two uncorrelated noise carriers", provides a reference against which to compare the neuronal responses elicited by lead/lag pairs. Specifically, the superposition of the leading and lagging sounds in each ear would be expected to cause the frequency-specific binaural cues at any given instant and in any given spectral band to correspond to different locations in auditory space, resulting in a *spatial* decorrelation. Given the sensitivity of space-map cells to these

frequency-specific binaural cues (Euston and Takahashi, 2002; Spezio and Takahashi, 2003), the effects of the lead/lag delay must be differentiated from the effects of spatial decorrelation. The neural responses to two uncorrelated noise carriers thus allow us to account for the effects of delay per se, independently of spatial decorrelation.

The responses of a typical neuron to single sources placed at target and

masker loci are shown as dot-raster displays for three envelope modulation depths (Figure 4B). The top and bottom halves of each panel show the responses to the target and masker, respectively. As shown, the target evoked a vigorous burst at the stimulus' onset that decayed to a lower steady-state level. This burst/steady-state response property is clearest in the 0% modulation condition (Figure 4B, left). The responses increased slightly with increased modulation depths (Figure 4B, center, right). By contrast, the masker, which was outside the cell's SRF, evoked few if any spikes regardless of modulation depth (Figure 4B, bottom rasters). Responses evoked in the same neuron by two uncorrelated noises are shown in the spike-raster displays of Figure 4C.

Responses evoked by lead/lag pairs having different combinations of OED and modulation depth are shown in Figure 4D. The top half of each panel shows the response when the target lagged, and the bottom half, when the target led. By and large, the firing rates are similar whether the target lagged or led, except in the case shown in the upper right, which depicts the response to a 3 ms OED and 100% modulation. In this case, the cell's response to a leading target is considerably stronger than its response to a lagging target. Finally, as with single sources and two uncorrelated noises, deeper modulations caused higher firing rates whether the target led or lagged.

The observations above are summarized for our sample population of cells in the bottom panels of Figure 4 (50 envelopes, one to three repetitions). Figure 4E plots against modulation depth the normalized firing rate of all of our cells evoked by single maskers (green), single targets (blue), and by two uncorrelated noises (dashed line). Overall, targets evoked strong responses that increased slightly with modulation depth. Maskers, by contrast, evoked little if any activity regardless of modulation

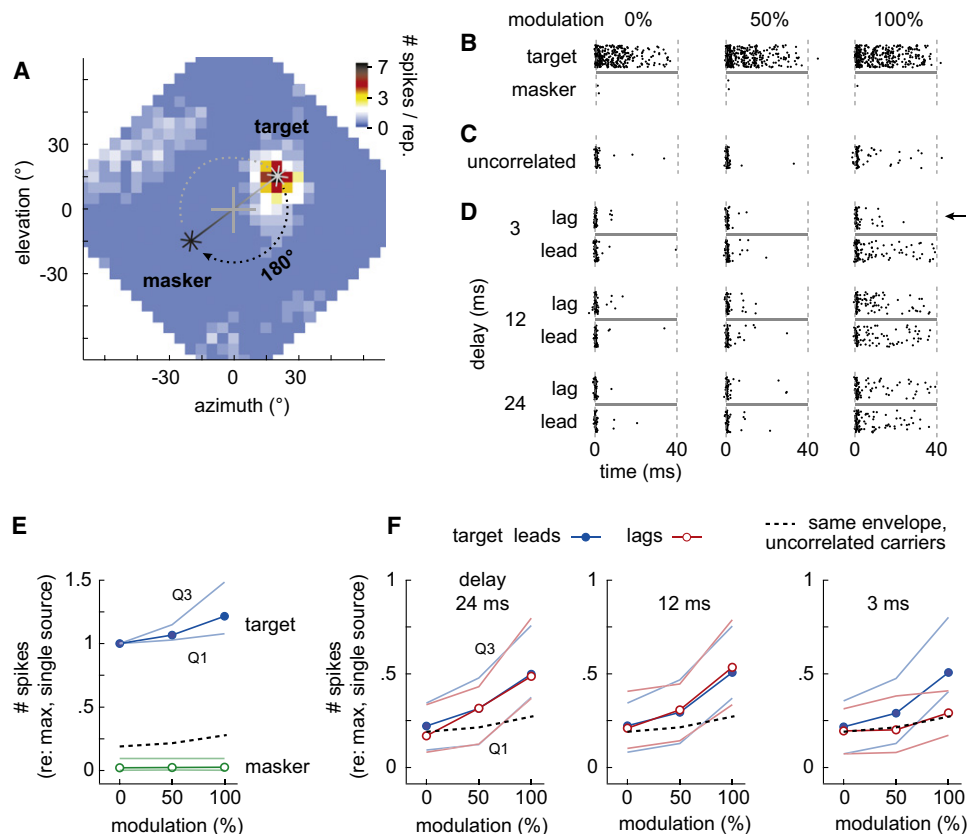


Figure 4. Neural Responses

(A) Spatial response of a typical space-map neuron. The octagonal region represents the locations in the frontal hemisphere from which virtual auditory stimuli were presented. Crosshairs at the center represent the center of the owl's gaze. Pixel colors represent the average spike count elicited by a single source emitting broadband noise (100 ms). This cell's spatial receptive field (SRF) was therefore in the upper-right quadrant of frontal space. In trials with lead/lag pairs, the source in the cell's SRF is referred to as the *target*, and the other source, which was always placed diametrically across from the center of gaze, is referred to as the *masker*.

(B) Spike-raster plots of neural responses elicited by the target (top row) and masker (second row) having envelope modulations of 0%, 50%, and 100%.

(C) Responses evoked by two *uncorrelated* noises (see text).

(D) Responses evoked by lead/lag pairs analogous to those used in the behavioral experiments (envelope modulations of 0%, 50%, 100%; OEDs of 3, 12, 24 ms). Responses evoked by lagging targets are shown in the rows above those evoked by leading targets.

(E) Median spike counts (normalized) plotted against modulation depth for targets (blue), maskers (green unfilled circles), and two uncorrelated noises (black dashed line). Thin lines demarcate the first and third quartiles about the median.

(F) Median spike counts (normalized) elicited by leading (blue filled circles) and lagging targets (red unfilled circles) and by two simultaneous sources (black dashed lines) are plotted against modulation depth for OEDs of 24 ms (left), 12 ms (center), and 3 ms (right).

depth. The responses to two uncorrelated noises (dashed line) fell between these two extremes.

Each panel in Figure 4F shows the responses of cells to lead/lag stimuli plotted against modulation depth for OEDs of 24 (left), 12, or 3 ms. Responses to leading targets (blue) increased with modulation depth but did not vary with OED ($p < 1 \times 10^{-6}$; $df = 26$; Kruskal-Wallis, Student-Newman-Keuls test, $p = 0.05$). Responses to the lagging targets (red) were similar to those of leading targets at the longer delays of 12 and 24 ms.

When the OED was 3 ms (Figure 4F, right), however, we observed an interesting difference in the responses to the lead and lag relative to those evoked by two uncorrelated noises. Specifically, the responses to the lead sound for the deeper modulations were stronger than those evoked by two uncorrelated

noises, while the responses to the lagging sound were similar to *but not less than* those evoked by two uncorrelated noises. In other words, the weak response to the lag simply reflects the effects of spatial decorrelation resulting from the superposition of two sounds. There is no evidence of a suppression of the lag response that is attributable to the delay per se.

Our behavioral and physiological experiments show that the conditions that result in localization dominance (Figures 2 and 3), i.e., short OEDs and deep modulations, are also those that result in a preferential neuronal response to the leading source (Figure 4). This correlation between the differential neural responses to the leading and lagging sources and localization dominance has been pointed out earlier, beginning with the work of Yin (1994). What causes the lead sound to evoke a

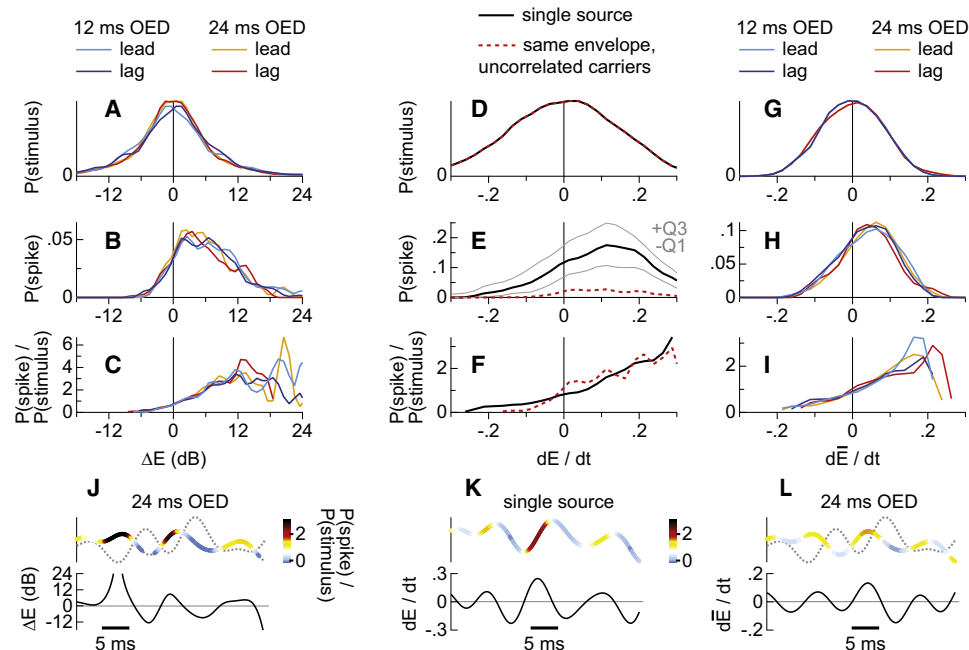


Figure 5. Generic Neuronal Preferences

(A) Distribution of ΔE values obtained from lead/lag pairs with OEDs of 12 ms (blue lines) or 24 ms (red or orange lines) and 100% modulation. At these OED values, the carriers and envelopes are uncorrelated within a short time window (see text). Positive values indicate that the target's amplitude was higher than the masker's.

(B) Spiking probability plotted against ΔE shows that neurons discharged when the target's amplitude was higher than the masker's. The ordinate represents median spike probability per stimulus and sampled cell.

(C) Spiking probability scaled to the likelihood of ΔE values, obtained by dividing the distributions in (A) and (B).

(D–F) Distributions analogous to (A)–(C) for dE/dt , measured for single sounds and two uncorrelated noise carriers (dashed line). Positive derivatives indicate that the envelope's amplitude was rising. Thin lines in (E) show the first and third spike-probability quartiles for the sampled cells ($n=28$).

(G–I) Distributions analogous to (A)–(C) for $d\bar{E}/dt$ and lead/lag pairs having OEDs of 12 and 24 ms (solid lines). Similarity of (F) and (I) suggest that the neurons' preference for positive derivatives remains similar whether there are two sounds or just one.

Upper traces in (J)–(L) show representative lead/lag envelope pairs; the thicker, colored line represents the target's waveform while the dotted line represents the masker's waveform. The color of the target's envelope indicates spike probability (warmer colors represent higher probabilities). Lower traces show ΔE (J), dE/dt (K), and $d\bar{E}/dt$ (L) from which the probabilities were estimated.

See also Figure S2.

greater neural response than that evoked by the lag sound at short delays, or by two uncorrelated noises, and why does this effect depend upon modulation depth?

Analyses of Acoustical Cues and Neuronal Responses

We hypothesize that the stronger neural response to the leading sound (above) might be explained by a combination of simple acoustical principles and the firing properties of space-map neurons.

First, neurons in the owl's space map are highly selective for the binaural cues that correspond to the locations of their SRFs (Figure 4A). When two sources concurrently emit sounds with overlapping amplitude spectra, the frequency-specific binaural cues assume values that are the average of each individual source's cues weighted by their amplitudes. Thus, if one of the two sources momentarily achieves a higher amplitude than the other, the binaural cues approximate the values of the higher-amplitude source for that moment (Blauert, 1997; Keller and Takahashi, 2005; Snow, 1954; Takahashi and Keller, 1994). A space-map neuron is therefore likely to fire when the

amplitude modulation at a given instant favors the target, which is in its SRF, over the masker. This preference for amplitudes that favor the target is shown in Figures 5A–5C. Figure 5A shows the distribution of the relative amplitudes of target and masker, ΔE , sampled at regular (0.1 ms) intervals for conditions in which the two sounds were effectively uncorrelated (OEDs of 12 and 24 ms). The distribution in Figure 5A shows that positive (target > masker) and negative (target < masker) ΔE values were equally likely. The relative amplitude was also sampled within a ± 1 ms time window surrounding each spike, instead of at regular intervals, and accumulated into the distribution of Figure 5B (28 cells). This distribution, by contrast, demonstrates that the probability of spiking is highest when $\Delta E > 0$. Figure 5C scales the spiking probability by the frequency of occurrence of ΔE .

Second, neurons of the owl's space map, like those of many other auditory structures, are especially sensitive to the onsets of sounds. When a sound's envelope is unmodulated, the only onset is at the beginning of the sound. The firing rate is correspondingly highest there and diminishes as the sound

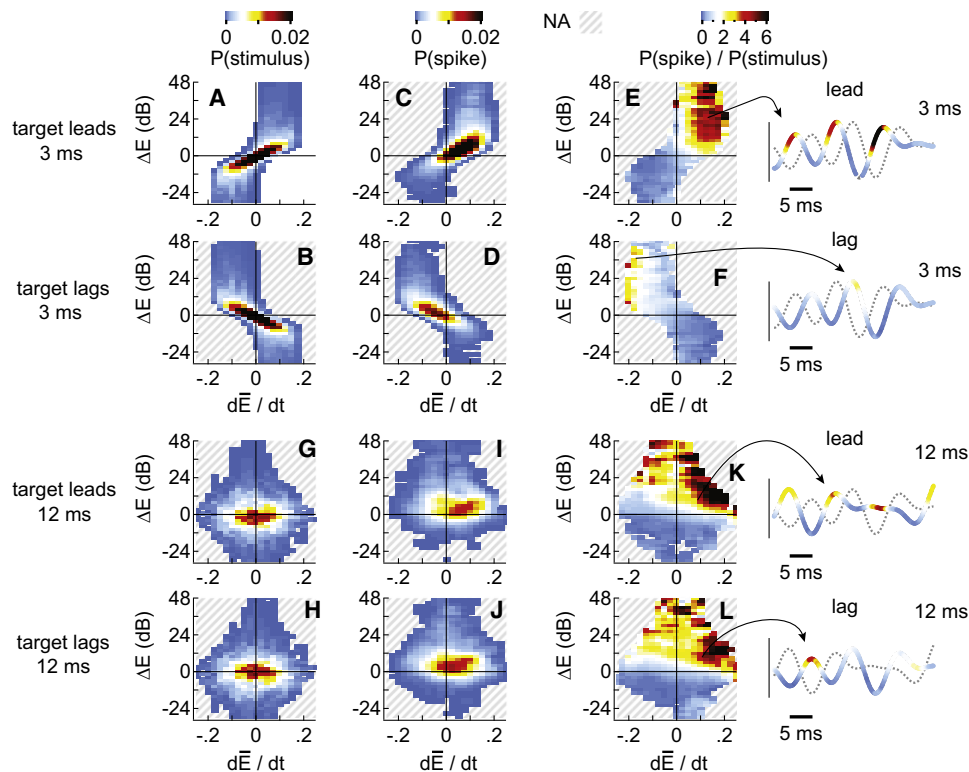


Figure 6. Analysis of Spike-Conductive Features (100% Modulation)

To analyze the stimulus envelopes, ΔE and $d\bar{E}/dt$ (see text) were measured at regular intervals for the 50 envelope pairs used in our experiments. Each measurement was then assigned to a frequency-distribution histogram with axes of $d\bar{E}/dt$ and ΔE . Distributions are shown for four stimulus conditions: 3 ms OED (A and B) or 12 ms OED (G and H); target leading (A and G) or lagging (B and H). Analogous measurements were then made at the times when spikes were evoked (C, D, I, and J). Values represent median spike probability, per histogram bin, stimulus, and sampled cell. Cross-hatches represent areas for which there were no values. After scaling each stimulus distribution to the volume under each spike-specific distribution, spike-specific distributions were divided by the stimulus distributions. These quotients are shown for the same stimuli in (E), (F), (K), and (L). To the right of each panel are representative lead/lag envelope pairs; the thicker, colored line represents the target's waveform, while dotted lines represent the masker's waveform. The color of the target's envelope indicates spike probability (warmer colors represent higher probabilities).

continues (Figure 4B; 0% depth). For the more deeply amplitude-modulated sounds, the stimulus resembles a series of distinct peaks, and the rising edge of each such peak may effectively serve as an onset-like event. A space-map neuron is therefore likely to fire on the rising edges of single sounds or when the target's envelope, which is in its SRF, is rising. The preference for rising amplitudes is shown in Figures 5D–5F, which are analogous to Figures 5A–5C, except that the stimuli analyzed were the envelopes of single sounds or two uncorrelated noises. As shown (Figures 5E and 5F), the probability of spiking is highest when the derivative of an envelope segment (dE/dt) is positive even though positive and negative derivatives were equally likely (Figure 5D). Figures 5G–5I show the distributions related to $d\bar{E}/dt$, but for pairs of sounds. For this analysis, the derivative $d\bar{E}/dt$ was measured using the composite envelope \bar{E} of the masker and target ($[E_{\text{target}} + E_{\text{masker}}]/2$), as shown for a representative envelope pair in Figure 5L. Importantly, the distributions of spikes (Figures 5H and 5I) show a bias for positive values much like those for single sources (Figures 5E and 5F), suggesting that a cell's preference for rising amplitudes remains similar in one- and two-source

conditions. Additional distributions for an OED of 3 ms and 50% modulations are shown in Figure S2.

We next tested our hypothesis (above) that it is the *conjunction* of these two stimulus features, ΔE and $d\bar{E}/dt$, that determines the responses of space-map neurons to lead/lag pairs. More specifically, space map neurons are most likely to discharge when the amplitude of the target exceeds that of the masker ($\Delta E > 0$) while the composite envelope is rising ($d\bar{E}/dt > 0$). According to our hypothesis, the leading source elicits a higher spike rate because these spike-conductive conditions are more abundant for the lead than for the lag when OEDs are short and modulations are deep. With longer OEDs and shallower modulations, the advantage for the leading sound disappears. To test this hypothesis, estimates of the derivative and difference were assigned to (three dimensional) frequency distributions having axes of ΔE and $d\bar{E}/dt$. This procedure was repeated for all 50 pairs of envelopes used in the experiments.

Figures 6A and 6B show, respectively, the histograms for leading and lagging sounds that were 100% modulated and presented with an OED of 3 ms. Warmer colors represent high incidences of the derivatives and differences. As shown, there

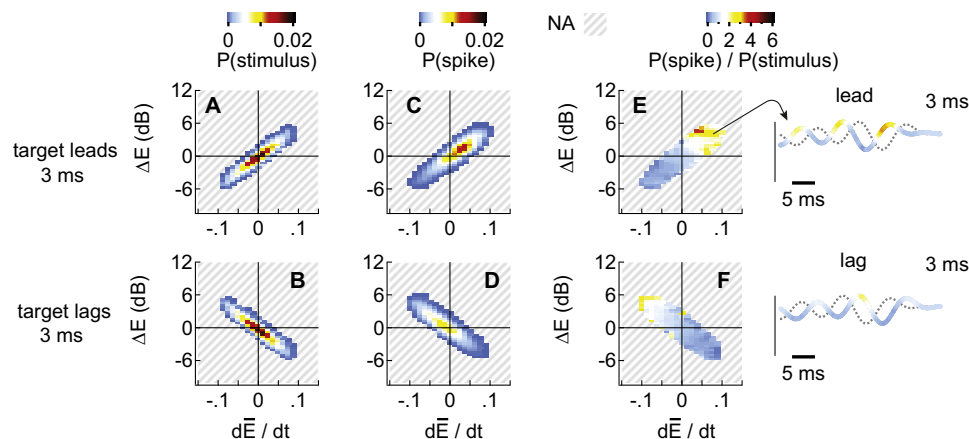


Figure 7. Analysis of Spike-Conductive Features (50% Modulation)

Plots shown here are analogous to those in Figures 6A–6F except that $d\bar{E}/dt$ and ΔE were measured when the stimuli were modulated at 50% (OED = 3 ms).

is a peak near the origin for both the leading (Figure 6A) and lagging (Figure 6B) conditions. However, it is also clear that for the leading sounds, there are more events in the upper right quadrant, which represents the spike-conducive conditions ($d\bar{E}/dt > 0$; $\Delta E > 0$; Figure 5) that are rare in the histogram for the lagging sounds. This analysis suggests that leading sounds evoke more spikes simply because the conjunction of conditions that favor spiking are more prevalent in the leading sound than in the lagging sound for an OED of 3 ms and 100% modulation.

This difference in the statistics of the leading and lagging envelopes is accentuated by the space-map neurons that are highly sensitive to the spike-conductive conditions. To quantify this sensitivity, we first estimated ΔE and dE/dt values within a ± 1 ms time window surrounding each spike (instead of at regular intervals) and assigned them to a frequency distribution. The spike-specific distributions are shown in [Figures 6C](#) and [6D](#) for leading and lagging targets, respectively, normalized to the volumes under the distributions. As one might expect, there are more events for the leading sound than for the lagging sound because leading sounds evoke more spikes. To determine the efficacy of the spike-conductive conditions, we divided the normalized spike-specific distributions ([Figures 6C](#) and [6D](#)) by the stimulus distributions ([Figures 6A](#) and [6B](#)). The quotients, shown in [Figures 6E](#) and [6F](#) for the leading and lagging targets, respectively, are quite different, with far more events in the upper right quadrant for the lead ([Figure 6E](#)) than the lag ([Figure 6F](#)). Inspection of the quotient histogram suggests that the cells are very sensitive to positive values of ΔE and dE/dt , especially when they are large. These large values are absent in the lagging sound's envelopes.

Spike probabilities computed from the quotient plots (Figures 6E and 6F) are superimposed on examples of the leading and lagging envelopes in the rightmost column of Figure 6. As shown in the upper two pairs of envelopes, spiking probabilities are higher for the lead than lag at 3 ms and occur along the rising edges of the lead envelope (warmer colors) and when the leading envelope has a higher amplitude than the lagging envelope.

Similar analyses were conducted for longer OEDs and shallower modulations, i.e., the conditions under which localization

dominance failed. [Figures 6G and 6H](#) show the stimulus distributions for the leading and lagging sounds at an OED of 12 ms and a 100% modulation depth. Similar results were obtained with an OED of 24 ms and are not shown. Comparison of the two distributions shows that both the leading and lagging stimuli have an abundance of events in the upper right quadrant where the conjunction of spike-conductive values are represented, and importantly, there is little lead/lag asymmetry. Correspondingly, neither the spike-specific distribution ([Figures 6I and 6J](#)), nor their quotients ([Figures 6K and 6L](#)) show a lead/lag asymmetry. The abundance of spike-conductive events for the leading and lagging stimuli is consistent, however, with the observation that both the leading and lagging stimuli evoke a high response ([Figure 4](#)). Thus, the responses to both the lead and lag stimuli are strong and equal.

Figures 7A and 7B show the stimulus distributions for the leading and lagging sounds at an OED of 3 ms and a 50% modulation depth. As was the case when the stimuli were deeply modulated (Figures 6A–6F), the stimulus distributions for the leading and lagging sounds are asymmetric, i.e., there are more spike-conducive events ($d\bar{E}/dt > 0$; $\Delta E > 0$) for the leading sound than for the lagging sound. However, when the modulations are shallow and the amplitudes change slowly, neither the leading nor the lagging sound has a significant advantage over the other (Figures 7C–7F) because the magnitudes of $d\bar{E}/dt$ and especially ΔE remain small. The small magnitudes are, in turn, consistent with the weak responses to leading and lagging stimuli with 50% modulation (Figure 4).

Model Verification

As an indication of how these spike-conductive conditions might account for the asymmetry in responses to lead and lag sounds, we parameterized a model of the average space-map neuron with the sensitivity to $d\bar{E}/dt$ and ΔE derived from quotient histograms (e.g., [Figures 6K and 6L](#)) obtained with long OEDs (12 and 24 ms). We then predicted the responses evoked by an OED of 3 ms.

Quotients corresponding to the longer delays were first averaged (Figure 8A). Then, for every envelope in each of the

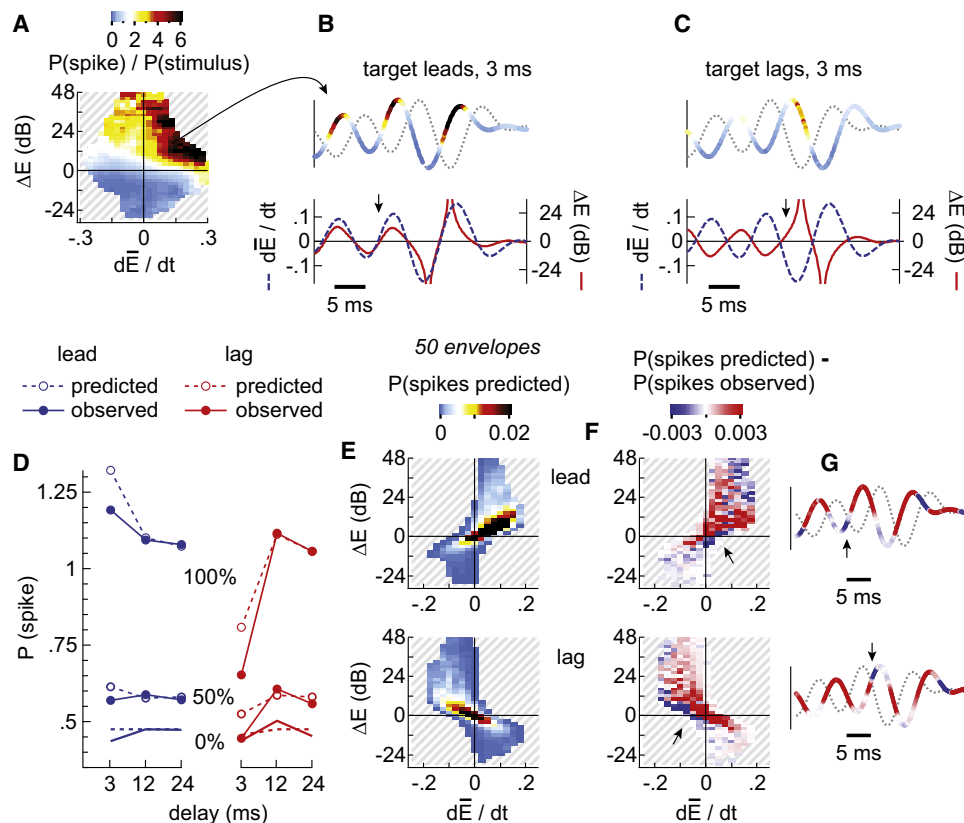


Figure 8. Model Verification

(A) The average preferences of a space-map neuron for combinations of ΔE and $d\bar{E}/dt$ were derived from the quotient histograms obtained with OEDs of 12 and 24 ms, 100% modulation (see also Figures 6K and 6L) and used to predict responses obtained for stimuli with 3 ms OEDs and other modulation depths. The upper traces in (B) and (C) show representative lead (B) and lag (C) target envelopes color coded according to their spike probabilities; derived from the quotient histogram (A). The lower traces show the derivative (blue) and differences (red) for the upper traces.
 (D) Overall predicted (open circles) and observed (filled circles) spike probabilities for lead (blue) and lag (red) targets. Not surprisingly, the predictions are nearly perfect for the 12 and 24 ms OEDs, the values on which the model neuron was parameterized.
 (E) Spike probabilities for lead and lag targets assigned to two-dimensional distributions.
 (F) Spikes predicted (E) minus spikes observed (e.g., Figures 6C and 6D) per histogram bin, stimulus, and sampled cell.
 (G) Representative lead and lag target envelopes color coded according to the differences between spikes predicted and spikes observed (F). Bluish colors and arrows indicate where the neural responses were underestimated.

50 lead/lag pairs used in the physiology experiments (3 ms OED), we estimated the probability of a spike given the model neuron's sensitivity to ΔE and $d\bar{E}/dt$ values. Estimates for representative lead and lag targets are shown in Figures 8B and 8C, respectively. The upper traces in each panel represent a lead/lag pair; the thicker, colored line represents the target while the dotted line represents the masker. The coloration of the envelope predicts the probability that the envelope segment would result in a spike, given its ΔE and $d\bar{E}/dt$ values. Note that spike-conducive conditions (warmer colors) were predicted to arise more frequently, and for longer periods of time, when the target led than when it lagged. Correspondingly, the lower traces show that both positive differences and derivatives coincided when the target led (Figure 8B) but were out of phase when the target lagged (Figure 8C). The spiking probabilities then were accumulated into distributions (Figure 8E) and normalized to the overall probability with which spikes were evoked when the OED was

long and the target led or lagged (i.e., to the volumes under plots shown in Figures 6K and 6L).

Figure 8D compares the predicted (dashed lines) and observed (solid lines) spike probabilities for each of the modulation depths and OEDs. Because the model was parameterized using data when the OED was 12 and 24 ms, predictions for these stimuli were nearly perfect. The model was also reasonably accurate, however, in predicting the overall probabilities when the OED was 3 ms, although it slightly overestimated the responses to both the leading and lagging sounds.

Figure 8F, which plots the difference between predicted (Figure 8E) and observed spike probabilities (Figures 6C and 6D), confirms that spike probabilities were lower than predicted for large areas of feature spaces (red areas). A subtle but potentially interesting exception is indicated by the arrows pointing to a blue region near the origins of the upper (target led) and lower plots (target lagged). Projecting these spike probabilities onto the

traces of the representative waveforms (Figure 8G), we see that the cells fired more spikes than predicted immediately after a peak in the masker's envelope, whether the target led (top, blue segments, arrows) or lagged (bottom). Quotients used to predict spikes were consistently small near the origin (Figure 8A), and the differences shown (Figures 8F and 8G) should therefore be viewed with caution. However, one might have expected even fewer spikes after a peak in the masker's envelope (i.e., the source outside the SRF) if the representations of the two sound sources on the map were inhibiting one another.

DISCUSSION

Owls, like humans (Dizon and Colburn, 2006; Zurek, 1980), behaved in a manner consistent with localization dominance with stimulus pairs lacking onset (and offset) delays, provided that the envelopes were well modulated. Neurons in the owl's space map also fired preferentially to the leading stimuli at short delays, but again, only when the modulations were deep. Analysis of the signals suggests that the acoustical conditions that favor spiking in the space map are more abundant for the leading sound at short delays than for the lag. At long delays or with shallow modulations, this asymmetry in the leading and lagging signals disappears, as does localization dominance and its neural parallel.

Integration of Acoustics, Physiology, and Behavior

Our results suggest that localization dominance in the owl is a byproduct of generic neuronal preferences for onsets and binaural cues. Thus, when modulations are deep and the delay is short, an envelope peak in the lead waveform is likely to be present by itself, briefly, causing the frequency-specific binaural cues to approximate values at the lead source's location. Moreover, its rising edge is unobstructed by the corresponding peak in the lagging waveform. The space-map neurons therefore respond vigorously, as they would to one of a series of brief transients from a single (i.e., unmasked) source. The requirement for a rising edge may be related to the finding that human localization dominance is absent if a signal's rise time is slow (Rakerd and Hartmann, 1985). The corresponding peak in the lag and its rising edge, on the other hand, are likely to be superimposed with its counterpart in the lead, resulting in a spatial decorrelation. The space-map neurons, being sensitive to the coherency of spatial cues (Albeck and Konishi, 1995; Euston and Takahashi, 2002; Spezio and Takahashi, 2003; Yin et al., 1987), would therefore respond weakly.

Evidence for localization dominance and its neural parallel were not observed at longer delays or with shallower modulations. At the longer delays, the peaks in the lag and their rising edges were no longer masked by the corresponding peaks in the lead, evoking responses similar to those of the lead. With shallower modulations, neither source had an amplitude advantage over the other ($\Delta E \approx 0$), and the rises in amplitudes were slow ($dE/dt \approx 0$). Thus, for delays tested here and a short integration time window (Joris et al., 2006; Siveke et al., 2008; Wagner, 1991, 1992), the lead/lag pair becomes equivalent to two, simultaneous, uncorrelated noises, to each of which, the space map neurons respond weakly (Keller and Takahashi, 2005).

In nature, reflections are not simply delayed copies of the direct sound but have lower overall amplitudes. In terms of our model, the echo's weakness would mean that ΔE would favor the direct sound consistently, thus further strengthening localization dominance.

Principles Shared with Other Models

Our explanation is related to that of Tollin (1998) and Trahiotis and Hartung (2002) who attributed the weak response of an IC neuron to a lagging click (Fitzpatrick et al., 1995; Litovsky and Yin, 1998; Yin, 1994) to the generation of binaural cues that do not correspond to the neuron's preferred values. These suboptimal binaural cues were attributed, in turn, to the superposition of the ringing responses of cochlear filters, passed through a hair cell model (Meddis et al., 1990), evoked by the lead and lag clicks in each ear (Hartung and Trahiotis, 2001). In our "envelope" hypothesis and in this "ringing" hypothesis, it is the superposition of waveforms, whether internal (Hartung and Trahiotis, 2001; Trahiotis and Hartung, 2002) or external, that alters the binaural cues and weakens the responses of neurons to the lagging sound, not neural inhibition.

Our explanation is also reminiscent of the process of "glimpsing," in which a masked target sound is tracked by catching glimpses of the target during quiet periods in the masker (Miller and Licklider, 1950). With the stimuli used presently, glimpses of the lead or lag target become available whenever ΔE favors one source over the other. Thus, deep modulations tend to offer more such glimpses than shallow ones. The idea of glimpsing is also consistent with our recent study of unmodulated sounds (Nelson and Takahashi, 2008) which showed that the owl's tendency to localize the lead or lag source depends on the lengths of the onset and offset time differences (Figure 1A) during which clear glimpses of the lead or lag sounds were available. It is also found in the study by Devore et al. (2009), which showed that sounds are localized in reverberant rooms by "glimpsing" the spatial cues available at the onset of the sound when reflections from room-surfaces had yet to arrive. Here again, specialized mechanisms to prevent responses to reflections need not be invoked.

Localization Precision

Envelope modulation depth had an effect on saccade precision, in addition to its effect on the proportion of turns toward one source or the other. Interestingly, high saccade precision was observed under conditions that elicited robust firing on the space map. For instance, at a 3 ms delay and 100% modulation, the firing evoked by the leading sound was considerably stronger than that evoked by the lag or by two uncorrelated noises. Correspondingly, the saccades to the lead were not just more numerous; they were more precise. The saccades to the lag, when they occurred, had considerably greater error. When conditions were such that both the lead and lag sources evoked a strong response (12 or 24 ms OED, 100% modulation), the precision of saccades to both sources was equal and approximated that to single sources. Conversely, if both sources evoked a weak response, as when the modulations were shallow (0%), saccades to both sources had large errors.

Although a high spike count is associated with high precision for trials employing lead/lag pairs, it is interesting to note that the responses to paired stimuli were considerably lower than those evoked by single stimuli (Figure 4), even though the saccades to some of these paired sounds (100% modulation) were nearly as precise as those to single sources. One interpretation may be that the putative function relating saccade precision and spike count is steep and asymptotic. Thus, the response to single sources is larger than needed for near-optimal precision. Precisely how vigorous firing observed in single units is related to saccade precision is unknown, although it is consistent with the observation that low-amplitude stimuli, which evoke fewer spikes on the space map, also elicit saccades with low precision (Whitchurch and Takahashi, 2006). One possibility is that high rates of firing in single space-map units reflect binaural cues that are stable over the duration of a sound, as is the case when there is only a single source. With a single sound source, the frequency-specific binaural cues remain at values corresponding to the source's location and only the space-map cells tuned to that combination of cues will fire vigorously, resulting in a well-focused pattern of activity across the map. However, when there are multiple sources concurrently emitting sounds with overlapping spectra, as is the case when lead and lag sounds overlap, the binaural cues fluctuate over time, and the patterns of fluctuation may differ in each spectral band. As a result, the binaural cues assume values that correspond to different locations from moment to moment, briefly exciting neurons tuned to those various locations. The spatial pattern of activity on the map, i.e., the neural image, may therefore jitter, with no strong focus remaining at any one location for long. The owl's saccadic system may make use of the space map's image in a way that causes saccade precision to depend upon the spatial pattern of activity on the map.

EXPERIMENTAL PROCEDURES

Animals

Experiments were carried out in six captive-bred adult barn owls (*Tyto alba*) under a protocol approved by the Institutional Animal Care and Use Committee of the University of Oregon. The birds were held in our colony under a permit from the US Fish and Wildlife Service (SCCL-723257).

Behavioral Subjects

Three barn owls (C, S, and T), housed together in a single enclosure within our breeding colony, were hand-reared and trained to make head saccades toward single visual and auditory stimuli for a food reward. The owls' weights were maintained at ~90% of the free-feeding level during training and testing.

Behavioral Apparatus

Behavioral experiments were conducted in a double-walled anechoic chamber (Industrial Acoustics Co. IAC; 4.5 m × 3.9 m × 2.7 m). An owl was tethered to a 10 × 4 cm perch mounted atop a 1.15 m post in the chamber. A dispenser for food rewards was also attached to the post near the owl's talons. A custom-built, head-mounted, magnetic search-coil system (Rommel Labs) was used to measure the owl's head movements in azimuth and elevation (Spitzer and Takahashi, 2006; Whitchurch and Takahashi, 2006). The coils were calibrated before each session to ensure accuracy within ±1° after adjusting for small differences in the angles with which head posts were attached to each owl's skull.

Sounds were presented from ten dome tweeters (2.9 cm, Morel MDT-39), located 1.5 m from the bird. Speaker pairs were positioned with opposite polar

angles (Figure 2) relative to the location of a central fixation LED (2.9 mm, $\lambda = 568$ nm). Radii (r) were standardized across speaker pairs at 10°, 15°, 20°, 25°, and 30°. New polar angles were generated randomly for each speaker pair every two to four test sessions. Subjects were monitored continuously throughout test sessions using an infrared camera and infrared light source (Canon Ci-20R and IR-20W).

Behavioral Stimuli

A session consisted primarily of single-source trials (~80%), amongst which we randomly interspersed paired-source trials. In paired-source trials, one of 50 different broadband (2–11 kHz) noises was randomly chosen and multiplied with one of 50 different low-pass-filtered (<150 Hz) envelopes (Figure 1C; 43, 52, or 64 ms initial durations). On a third of the trials, this envelope was flat and did not alter the natural envelopes of the noise carrier. On remaining trials, the envelope was scaled so that its deepest trough decreased to either 50% or 100% of its maximum (Figure 1B). Each modulated noise was then copied and the lag sound was delayed 3, 12, or 24 ms. Lead and lag sounds were 100% correlated at this time. Finally, both sounds were multiplied with a single 40 ms window (2.5 ms linear ramps) to synchronize their onsets and offsets (Figure 1A). Stimuli in single-source trials were drawn randomly from the pool of stimuli used in paired-source trials.

Stimuli were converted to analog waveforms at 48.8 kHz and routed to speakers using two power multiplexers (RP2.1, HB7, PM2R, Tucker-Davis Technologies). Stimulus sound-pressure level (SPL) was roved in 1 dB increments between 27 and 33 dB across trials, in reference to the output of an acoustic calibrator (re: 20 μ Pa, 4321, Brüel & Kjær). The owl's hearing threshold, measured with broadband noise, is approximately –9 dB SPL (Whitchurch and Takahashi, 2006). Stimuli were presented with equal SPL on paired-source trials.

Behavioral Paradigm

Head orientation and angular velocity were monitored continuously (20 Hz, RP2, Tucker-Davis Technologies) throughout test sessions. Stimuli were presented only after the owl fixated the centrally located LED for a random period of from 0.5 to 1 s. The LED was extinguished as soon as subjects oriented to within 3° of the LED. Trials were aborted if the owl moved beyond this 3° radius or if head velocity exceeded 2.6°/s before stimuli were presented. Once a trial was initiated, measurements were sampled at a rate of 1 kHz for 4.5 s, starting 1 s before the stimulus' onset. Subjects were automatically fed a small piece of a mouse (~1 g), after a ~1 to 3 s delay, if saccade velocity decreased to less than 4°/s and if the saccade ended at a location that was within 5°–8° from a previously activated speaker or member of a speaker pair.

Behavioral Data Analysis

After measuring baseline head velocity for 250 ms prior to stimulus presentation, the beginning of a saccade was defined as the time at which velocity exceeded this baseline measure by 5 SD for 50 ms. The end of a saccade was defined as the time at which velocity decreased to below baseline by 8 SD for 50 ms.

Each saccade's polar angle (α) was measured starting from the location where the saccade began to each subsequent point that was sampled along its trajectory (α_s), weighted by head velocity (v):

$$\alpha = \left(\sum \alpha_s \cdot v^3 \right) / \sum v^3 \quad (1)$$

Head-velocity was cubed because fewer points were sampled when head-velocity was high. Trajectories were measured in this way because the birds would frequently alter their trajectory as velocity decreased near the end of the saccade. Each saccade's radius was measured from its start to its end and assumed a linear trajectory.

Because the birds were first required to fixate a central LED, and because each speaker-pair was separated by ~180°, it was easy to determine whether saccades were lead or lag directed. Localization dominance was therefore estimated as the proportion of trials on which saccades were lag directed. Saccade angles were rarely >75° (absolute polar coordinates re: either source at 0°; see Figure 2; <1% of trials/subject) and were never equal with respect to

each paired source ($\pm 90^\circ$). In contrast with a similar previous study (Spitzer and Takahashi, 2006), saccades were rarely directed first toward one speaker and then toward the second, after making an $\sim 180^\circ$, hairpin, turn ($< 0.5\%$ of trials/subject). Reversals in saccade direction that did occur were as likely to occur on single-source trials and almost always corresponded to a spurious head movement prior to stimulus presentation.

Saccades with radii less than 5° or more than 45° were excluded from analysis. Also excluded were saccades to specific locations unassociated with the speakers (e.g., to the feeder or chamber door). Subject C rarely turned upward ($-90^\circ < \text{polar angle} < 90^\circ$; Figure 3A) when single sounds were presented and instead made highly stereotyped saccades to the left ($-90^\circ \pm 10^\circ$), right ($90^\circ \pm 10^\circ$), or to the feeder ($\sim 180^\circ$). These stereotyped saccades were excluded from analysis whether a single or paired stimulus was presented.

Neurophysiology

Experiments were conducted in a single-walled sound-attenuating chamber (Industrial Acoustics Co.; 1.8 m^3) equipped with a stereotaxic frame and surgical microscope.

Single-unit recordings were obtained from the space maps of three birds anesthetized with a mixture of ketamine and diazepam. Anesthesia was induced by an intramuscular injection of ketamine (22 mg/kg) and diazepam (5.6 mg/kg) and maintained with additional doses as needed (typically, every 2–3 hr after induction).

Upon induction of anesthesia, the owl was given a subcutaneous injection of physiological saline (10 cc; subcutaneous). Electrocardiographic leads and a temperature probe were attached and the owl was wrapped in a water-circulating blanket and placed in a stereotaxic frame. Axial body temperature was maintained between 35° and 40° C . Its head was secured to the frame by a previously implanted head post. Earphones were inserted into the ear canals, and a recording-well, implanted previously (Euston and Takahashi, 2002), was opened and cleaned to admit a glass-coated tungsten electrode ($1 \text{ M}\Omega < \text{impedance} < 12 \text{ M}\Omega$ at 1 kHz). The electrode's output was amplified and then fed to a signal processor (Tucker-Davis Technologies, RP2), oscilloscope, and audio monitor.

Typical recording sessions involved one to three electrode penetrations and lasted less than 12 hr. At the end of a session, the well was rinsed with a 0.25% solution of chlorhexidine and sealed. The owl was placed in an isolated recovery chamber until it recovered from anesthesia, after which, it was returned to its flight enclosure.

Neurophysiological Stimulus Presentation

Stimuli were presented in virtual auditory space (VAS) generated using individualized head related transfer functions (HRTF; Keller et al., 1998). HRTFs for each ear were measured from 613 locations in the frontal hemifield (Figure 1F), band-pass filtered, and stored digitally as 255 point (8.5 ms; 30 kHz) finite impulse response filters (HRIR). Stimuli were convolved with the HRIRs, converted to analog waveforms at 48.8 kHz, amplified (RP2.1, HB6, Tucker-Davis Technologies), and presented over the earphones (Etymotic Research, ER-1). Stimuli were presented 30 dB above the response threshold of each space-specific neuron at its best location.

We characterized each isolated unit's SRF by presenting sounds from 290 virtual locations. The "best location" within a SRF was defined as the location that evoked the maximum average spike count per repetition. Stimuli presented from this location in subsequent tests are referred to as *targets*. Stimuli presented from outside the SRF are referred to as *maskers*.

Stimuli were the same as those used in our behavioral experiments except that they were presented in VAS rather than in the free-field. For each cell, we also presented a pair of independent (i.e., 0% correlated) noise carriers simultaneously and without a delay from both the *masker* and *target* loci, as well as noise bursts from a single source at the location of the *masker* or the *target*. The stimuli ($n = 1050$) were presented in a random order over one to three repetitions. The interstimulus interval was 600 ms.

Neurophysiological Data Analysis

Spike times were adjusted by subtracting the median first-spike latency of each cell's response (9.4–14.3 ms; 0% modulation; 50–150 repetitions). Subtracting the median first-spike latency of all the cells (12.2 ms), as opposed

to each individual cell's measured latency, weakened but did not eliminate the overall biases that we observed for rising envelopes ($d\bar{E}/dt > 0$) and binaural cues ($\Delta E > 0$).

Each cell's response was quantified by counting the spikes that the *target* evoked within a 50 ms time window (-5 to 45 ms; Figures 4B and 4C). The spike count was then normalized to that evoked, in each cell, by single 40 ms sounds from the center of its SRF (Yin, 1994). We used nonparametric statistics because variation was often skewed in a positive direction when responses were weak.

To quantify the times of occurrence of spikes relative to the amplitude modulations, stimuli and neural responses were further analyzed by measuring (1) the instantaneous amplitude of each target's envelope in relation to the masker's ($\Delta E = \text{target dB} - \text{masker dB}$) and (2) the rate with which the composite envelope ($\bar{E} = E_{\text{target}} + E_{\text{masker}} / 2$) was increasing or decreasing ($d\bar{E}/dt$, $t = \pm 1$ ms). Differences and derivatives were sampled regularly (0.1 ms interval) across each of the 50 envelopes or were sampled at the times when spikes were evoked. Times near the onsets and offsets of the sounds (± 5 ms) were excluded to avoid sampling each stimulus' onset and offset ramp.

SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures and can be found with this article online at doi:10.1016/j.neuron.2010.07.014.

ACKNOWLEDGEMENTS

We thank Drs. C.H. Keller and E.A. Whitchurch for technical support and helpful discussions. This study was supported by grants from the National Institutes of Deafness and Communication Disorders: (F32) DC008267 to B.S.N. and (RO1) DC003925.

Accepted: June 22, 2010

Published: August 25, 2010

REFERENCES

- Albeck, Y., and Konishi, M. (1995). Responses of neurons in the auditory pathway of the barn owl to partially correlated binaural signals. *J. Neurophysiol.* 74, 1689–1700.
- Bala, A.D.S., Spitzer, M.W., and Takahashi, T.T. (2003). Prediction of auditory spatial acuity from neural images on the owl's auditory space map. *Nature* 424, 771–774.
- Bala, A.D.S., Spitzer, M.W., and Takahashi, T.T. (2007). Auditory spatial acuity approximates the resolving power of space-specific neurons. *PLoS ONE*, e675.
- Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization*, Revised Edition (Cambridge, MA: The MIT Press).
- Burger, R.M., and Pollak, G.D. (2001). Reversible inactivation of the dorsal nucleus of the lateral lemniscus reveals its role in the processing of multiple sound sources in the inferior colliculus of bats. *J. Neurosci.* 21, 4830–4843.
- Dent, M.L., and Dooling, R.J. (2004). The precedence effect in three species of birds (*Melospittacus undulatus*, *Serinus canaria*, and *Taeniopygia guttata*). *J. Comp. Psychol.* 118, 325–331.
- Dent, M.L., Klump, G.M., and Schwenzfeier, C. (2002). Temporal modulation transfer functions in the barn owl (*Tyto alba*). *J. Comp. Physiol. [A]* 187, 937–943.
- Devore, S., Ihlefeld, A., Hancock, K., Shinn-Cunningham, B., and Delgutte, B. (2009). Accurate sound localization in reverberant environments is mediated by robust encoding of spatial cues in the auditory midbrain. *Neuron* 62, 123–134.
- Dizon, R.M., and Colburn, H.S. (2006). The influence of spectral, temporal, and interaural stimulus variations on the precedence effect. *J. Acoust. Soc. Am.* 119, 2947–2964.

- Euston, D.R., and Takahashi, T.T. (2002). From spectrum to space: the contribution of level difference cues to spatial receptive fields in the barn owl inferior colliculus. *J. Neurosci.* 22, 264–293.
- Fitzpatrick, D.C., Kuwada, S., Batra, R., and Trahiotis, C. (1995). Neural responses to simple simulated echoes in the auditory brain stem of the unanesthetized rabbit. *J. Neurophysiol.* 74, 2469–2486.
- Haas, H. (1951). Über den einfluss eines einfachechos auf die harsamkeit von sprache. *Acustica* 1, 49–58.
- Hartung, K., and Trahiotis, C. (2001). Peripheral auditory processing and investigations of the “precedence effect” which utilize successive transient stimuli. *J. Acoust. Soc. Am.* 110, 1505–1513.
- Joris, P.X., Louage, D.H., Cardoen, L., and van der Heijden, M. (2006). Correlation index: a new metric to quantify temporal coding. *Hear. Res.* 216–217, 19–30.
- Keller, C.H., and Takahashi, T.T. (1996a). Binaural cross-correlation predicts the responses of neurons in the owl's auditory space map under conditions simulating summing localization. *J. Neurosci.* 16, 4300–4309.
- Keller, C.H., and Takahashi, T.T. (1996b). Responses to simulated echoes by neurons in the barn owl's auditory space map. *J. Comp. Physiol. [A]* 178, 499–512.
- Keller, C.H., and Takahashi, T.T. (2000). Representation of temporal features of complex sounds by the discharge patterns of neurons in the owl's inferior colliculus. *J. Neurophysiol.* 84, 2638–2650.
- Keller, C.H., and Takahashi, T.T. (2005). Localization and identification of concurrent sounds in the owl's auditory space map. *J. Neurosci.* 25, 10446–10461.
- Keller, C.H., Hartung, K., and Takahashi, T.T. (1998). Head-related transfer functions of the barn owl: measurement and neural responses. *Hear. Res.* 118, 13–34.
- Knudsen, E.I., and Konishi, M. (1978). Center-surround organization of auditory receptive fields in the owl. *Science* 202, 778–780.
- Lindemann, W. (1986). Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals. *J. Acoust. Soc. Am.* 80, 1608–1622.
- Litovsky, R.Y., and Yin, T.C.T. (1998). Physiological studies of the precedence effect in the inferior colliculus of the cat. II. Neural mechanisms. *J. Neurophysiol.* 80, 1302–1316.
- Litovsky, R.Y., Colburn, H.S., Yost, W.A., and Guzman, S.J. (1999). The precedence effect. *J. Acoust. Soc. Am.* 106, 1633–1654.
- Meddis, R., Hewitt, M.J., and Shackleton, T.M. (1990). Implementation details of a computational model of the inner hair-cell auditory-nerve synapse. *J. Acoust. Soc. Am.* 87, 1813–1816.
- Miller, G.A., and Licklider, J.C.R. (1950). The Intelligibility of Interrupted Speech. *J. Acoust. Soc. Am.* 22, 167–173.
- Nelson, B.S., and Takahashi, T.T. (2008). Independence of echo-threshold and echo-delay in the barn owl. *PLoS ONE*, e3598.
- Pecka, M., Zahn, T.P., Saunier-Rebori, B., Siveke, I., Felmy, F., Wiegrebe, L., Klug, A., Pollak, G.D., and Grothe, B. (2007). Inhibiting the inhibition: a neuronal network for sound localization in reverberant environments. *J. Neurosci.* 27, 1782–1790.
- Rakerd, B., and Hartmann, W.M. (1985). Localization of sound in rooms, II: The effects of a single reflecting surface. *J. Acoust. Soc. Am.* 78, 524–533.
- Siveke, I., Ewert, S.D., Grothe, B., and Wiegrebe, L. (2008). Psychophysical and physiological evidence for fast binaural processing. *J. Neurosci.* 28, 2043–2052.
- Snow, W.B. (1954). The effects of arrival time on stereophonic localization. *J. Acoust. Soc. Am.* 26, 1071–1074.
- Spezio, M.L., and Takahashi, T.T. (2003). Frequency-specific interaural level difference tuning predicts spatial response patterns of space-specific neurons in the barn owl inferior colliculus. *J. Neurosci.* 23, 4677–4688.
- Spitzer, M.W., and Takahashi, T.T. (2006). Sound localization by barn owls in a simulated echoic environment. *J. Neurophysiol.* 95, 3571–3584.
- Spitzer, M.W., Bala, A.D.S., and Takahashi, T.T. (2003). Auditory spatial discrimination by barn owls in simulated echoic conditions. *J. Acoust. Soc. Am.* 113, 1631–1645.
- Spitzer, M.W., Bala, A.D.S., and Takahashi, T.T. (2004). A neuronal correlate of the precedence effect is associated with spatial selectivity in the barn owl's auditory midbrain. *J. Neurophysiol.* 92, 2051–2070.
- Takahashi, T.T., and Keller, C.H. (1994). Representation of multiple sound sources in the owl's auditory space map. *J. Neurosci.* 14, 4780–4793.
- Tollin, D.J. (1998). Computational model of the lateralisation of clicks and their echoes. In *Proc. the NATO Advanced Study Institute on Computational Hearing*, S. Greenberg, M. Slaney, and I. Ciocco, eds. (NATO Scientific and Environmental Affairs Division), pp. 77–82.
- Trahiotis, C., and Hartung, K. (2002). Peripheral auditory processing, the precedence effect and responses of single units in the inferior colliculus. *Hear. Res.* 168, 55–59.
- Wagner, H. (1991). A temporal window for lateralization of interaural time difference by barn owls. *J. Comp. Physiol. [A]* 169, 281–289.
- Wagner, H. (1992). On the ability of neurons in the barn owl's inferior colliculus to sense brief appearances of interaural time difference. *J. Comp. Physiol. [A]* 170, 3–11.
- Wagner, H. (1993). Sound-localization deficits induced by lesions in the barn owl's auditory space map. *J. Neurosci.* 13, 371–386.
- Wallach, H., Newman, E.B., and Rosenzweig, M.R. (1949). The precedence effect in sound localization. *Am. J. Psychol.* 57, 315–336.
- Whitchurch, E.A., and Takahashi, T.T. (2006). Combined auditory and visual stimuli facilitate head saccades in the barn owl (*Tyto alba*). *J. Neurophysiol.* 96, 730–745.
- Yin, T.C. (1994). Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat. *J. Neurosci.* 14, 5170–5186.
- Yin, T.C.T., Chan, J.C.K., and Carney, L.H. (1987). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. *J. Neurophysiol.* 58, 562–583.
- Zurek, P.M. (1980). The Precedence effect and its possible role in the avoidance of interaural ambiguities. *J. Acoust. Soc. Am.* 67, 952–964.